Discovery report of Higgs Boson

Yixuan Chen, Harry Gardner, Hamzah Hashim, Vid Homšak, Anugi Wijayasiri

***Abstract*—This paper aims to identify the Higgs Boson by analysing a simulated rest mass spectra for the proton beam collisions. The background of the mass spetra follows an exponential distribution E ($A = 5.634 \pm 0.005 \times 10^4$, $\lambda = 30.02 \pm 0.04$), and the signal follows a Gaussian distribution G ($A = 6.23 \times 10^4$, $\mu = 125.2$, $\sigma = 1.34$). A χ^2 test was undertaken for both the background-only hypothesis and background plus signal hypothesis. Since for background-only hypothesis the p-value = 1.78$\times$10^{-8} = 5.51σ which is greater than significance level 5σ, and for background plus signal test the p-value = 0.41 which is much greater than the significance level, the background-only hypothesis is rejected and the existence of H is proved.**

I. Introduction

The Higgs field and its corresponding elementary particle, the Higgs Boson (H), was first theorised in 1964, offering a solution to the problematic unification of the weak interaction and electromagnetic interactions. The search, however, for proof of the H proved elusive until recent experiments at the Large Hadron Collier in 2012 announced the detection of a new particle thought to be the H [1,2].

This paper aims to simulate the results of the 2012 CMS experiment and to statistically determine whether the signal seen in the data from the experiment supports the existence, or lack thereof, of the H. By using the reduced χ_r^2 statistical distribution, we can determine if the data exhibits a statistically significant marker of the H.

II. Data Generation and Parameterisation

To determine whether the results of the LHC experiment prove statistically the existence of a new particle at the H's energy range, A simulated data set of the rest mass spectra for the proton beam collisions was created. It contains 9×10^5 background events, and 400 randomised events that model the H, which follow exponential and Gaussian distributions respectively.

From this, the data was plotted as a histogram, (see Fig. 1), using 30 bins in range of 104-155 GeV/c^2, along with the corresponding Poisson statistical uncertainties found by

$$\sigma_i = \sqrt{N_i} \tag{1}$$

The background distribution was first parameterised so that later the H could be isolated. To avoid H from influencing the background parameterisation, only events with rest mass values below 120 GeV were included, since theoretically $m_H = 125$ GeV/c^2. That yielded 71 bins on which, utilising scipy.optimize.curve - a non-linear least squares fit method, the exponential distribution E(x) was fitted on:

$$E(x) = Ae^{-\frac{x}{\lambda}}, \tag{2}$$

where $A = (5.634 \pm 0.005) \times 10^4$ is the amplitude and $\lambda = 30.02 \pm 0.04$ determines the gradient (see Fig. 1).

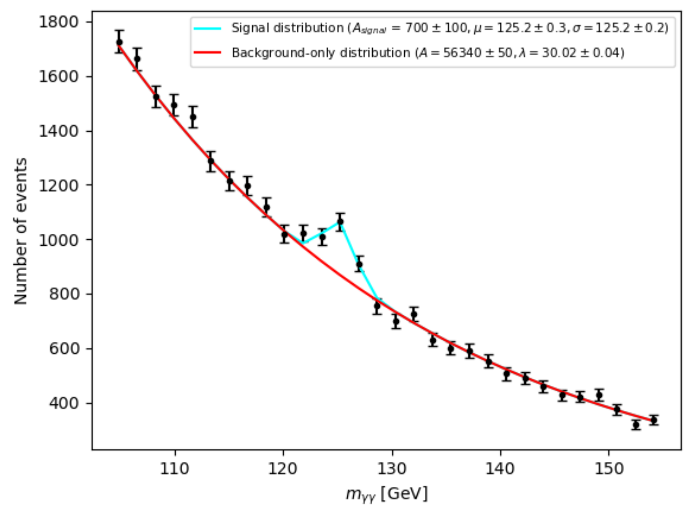

Fig. 1: **Background and signal hypothesis parameterisation.** Using scipy.optimize.curve exponential + Gaussian distributions were modelled with parameters: $A = 6.23 \times 10^4, \lambda = 29.35, \mu = 125.2$ and $\sigma = 1.34$

Furthermore, having parameterised the background, the Higgs signal could now be modelled, which follows Gaussian distribution $G(x)$ of the form

$$G(x) = \frac{A}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{3}$$

where, utilising scipy.optimize.curve, $A = 700 \pm 100$ is the amplitude, $\mu = 125.2 \pm 0.3$ is the mean of the signal and $\sigma = 125.2 \pm 0.2$ is the width of the Gaussian.

III. Hypothesis Testing

Having parameterised the background distribution, the goodness of the background fit could be tested. The value of $\chi_b^2 = 75.0$ was obtained and since number of degrees of freedom, $N_{dof} = 69$, the reduced $\chi_b^2 = \frac{\chi_b^2}{N_{dof}}$ was found to be 1.09. This suggests the background parameterisation is suitable.

Moreover, the hypothesis test was carried out. Taking background-only signal as the null-hypothesis H_0 modelled by E(x) in the range 104-155 GeV/c^2, where H should exist; for $N_{dof} = 28$, the value $\chi_b^2 = 90.3$ was found, which yields $\chi_{r\_b}^2 = 3.22$.

This means the function does not fit well. The p-value calculated using scipy.stats was found to be 1.78×10^{-8} significance level 5.51σ. This strongly suggests that H_0 could be rejected due to a poor goodness of fit, as the p-value is the probability of the null hypothesis H_0 being true given the hypothesis test.

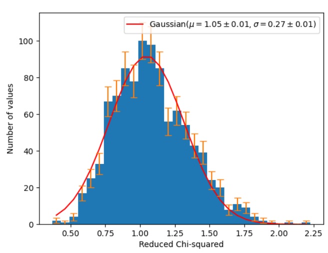

Fig. 2: **Distribution of the reduced χ_b^2** assuming H_0 in range 0-120 GeV/c^2.

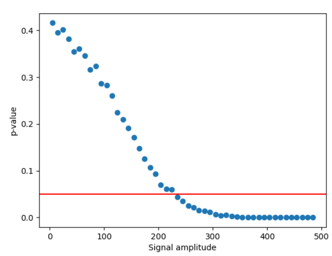

Fig. 3: **p-value as a function of signal amplitude.** The red line represents $p = 0.05$.

Furthermore, the simulation was repeated 1000 times to account for fluctuations in the background distribution, and hence the χ_r^2 distribution was obtained for the background-only fit (see Fig. 2). Since it is a random variable, it should be normally distributed; indeed best fit is given by $G(x; \mu = 1.05 \pm 0.01, \sigma = 0.27 \pm 0.01)$, suggesting that the goodness of the fit is very good.

Since p-value of χ_r^2 estimation of background signal varies with the signal amplitude, p-values were obtained by running the simulation for 500 values per interval and finding the average of the p-values generated in intervals of 10. The averages were plotted, as the values generated in the simulation were random and fluctuated around the lower signal amplitudes. The signal amplitude at which the expected p-value would equal 0.05 was found to be around 235. If a p value of less than 0.05 indicates a hint of a signal, at 0.05, the chance of finding a hint would be 0.95. This indicates it is highly possible that there is a signal of H in the data.

Then the χ_r^2 estimation of background plus signal hypothesis is tested. The estimation gave a reduced χ_r^2 value of $\chi_{r\ b+g}^2 = 1.03$ and a p-value of 0.41.

IV. Results and Analysis

Comparing $\chi_{r\_b+g}^2$ with that of background only hypothesis for the whole set of data, it is obvious that

$$\chi_{r\_b+g}^2 < \chi_{r\_b}^2, \tag{4}$$

Since the closer the χ_r^2 to 1, the better the estimation, the function includes the signal is a better estimation and it fits the data well.

If we set the level of significance of our hypothesis as 5σ, the p-value of background only estimation (1.78×10^{-8}) 5.51 σ away from the mean, which indicates the hypothesis that the signal does not exist should be rejected. In contrast, the p-value G(x) fitted on the signal is much larger than the level of significance, so the hypothesis that there is a Gaussian signal in the data is accepted.

To further show that the signal belongs to H's mass, the $\chi_{r\_b+g}^2$ over a range of masses is calculated (see Fig. 4). A significant drop of χ_r^2 can be observed around 125 GeV/c^2 which is the mass of H. This indicates the signal in data is due to H's mass, thus H do exist.

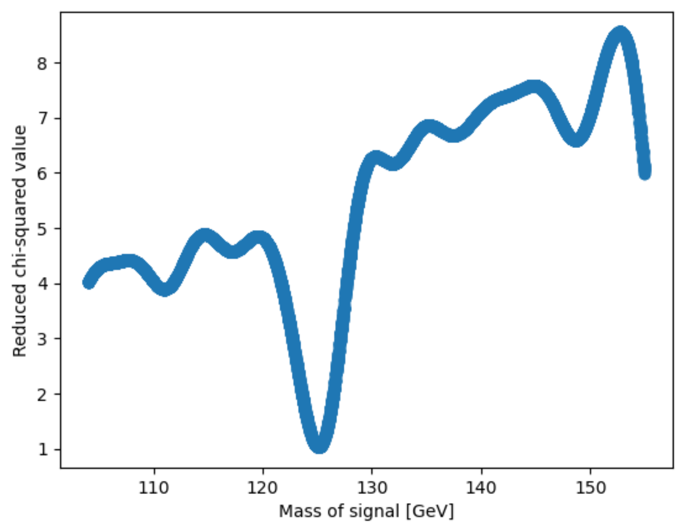

Fig. 4: **Reduced χ_r^2 as a function of signal mass position μ.** Reduced $\chi_r^2 \approx 1$ when $\mu = 125$

V. Conclusion

The theorized decay $H \rightarrow \gamma\gamma$ would generate a small irregular rest mass peak around 125Gev/c^2[4]. The background-only hypothesis test's parameterization is a poor fit, which signifies that the original background fit of the simulated data returns a p value of 1.78e-8 which is 5.51σ away from the mean which is greater than the significance level 5σ[4], hence the background-only hypothesis can be rejected. Varying the mass of supposed H, which follows normal distribution, the best fit of the signal was found at 125 GeV/c^2, which is coherent to the theorised rest mass model of the H providing strong evidence that the Higgs Boson exists.

References

[1] M. Tanabashi et al. (Particle Data Group) (2018). ”Review of Particle Physics”. Physical Review D. 98 (3): 030001. Bibcode:2018PhRvD..98c0001T. doi:10.1103/PhysRevD.98.030001

[2] Grey, H Mansoulie, B. (2020, May 30). The Higgs boson: The hunt, the discovery, the study and some future perspectives. Retrieved June 16, 2020, from https://atlas.cern/updates/atlas-feature/higgs-boson

[3] SToM Computing Problem - Discovering the Higgs Boson Dr Mark Richards